



Okasha, S. (2016). On Hamilton's Rule and Inclusive Fitness Theory with Non-additive Payoffs. *Philosophy of Science*, 83(5), 873-883.
<https://doi.org/10.1086/687871>

Publisher's PDF, also known as Version of record

Link to published version (if available):
[10.1086/687871](https://doi.org/10.1086/687871)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the final published version of the article (version of record). It first appeared online via University of Chicago Press at DOI: 10.1086/687871. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

On Hamilton's Rule and Inclusive Fitness Theory with Nonadditive Payoffs

Samir Okasha^{*†}

Hamilton's theory of inclusive fitness is a widely used framework for studying the evolution of social behavior, but controversy surrounds its status. Hamilton originally derived his famous $rb > c$ rule for the spread of a social gene by assuming additivity of costs and benefits. However, it has recently been argued that the additivity assumption can be dispensed with, so long as the $-c$ and b terms are suitably defined, as partial regression coefficients. I argue that this way of generalizing Hamilton's rule to the nonadditive case, while formally correct, faces conceptual problems.

1. Introduction. Hamilton (1964) derived his famous rule for the spread of an allele that causes a social behavior, $rb > c$, by assuming both additivity of costs and benefits and weak selection. Much subsequent work on inclusive fitness and kin selection has also relied on these assumptions, a fact that has led some biologists to complain that Hamilton's rule, and kin selection theory more generally, is of rather limited applicability (van Veelen 2009; Nowak, Tarnita, and Wilson 2010, 2011; Allen, Nowak, and Wilson 2013). In response to this complaint, recent studies by Gardner, West, and Wild (2011) and Marshall (2011, 2015) argue that when the ' r ', ' b ', and ' c ' terms are appropriately understood, Hamilton's rule is in fact a fully general and exact statement about natural selection that requires neither additivity nor weak selection. These studies build on earlier work in a similar vein by Queller (1992a, 1992b) and Frank (1997, 1998), who derived a generalized form of Hamilton's rule from the Price equation.

^{*}To contact the author, please write to: Department of Philosophy, 9 Woodland Rd., Bristol BS8 1TB, United Kingdom; e-mail: Samir.Okasha@bristol.ac.uk.

[†]Thanks to Johannes Martens, Jonathan Birch, Andy Gardner, James Marshall, and Alan Grafen for comments and discussion. This work was supported by the European Research Council Seventh Framework Program (FP7/20072013), ERC grant agreement 295449.

Philosophy of Science, 83 (December 2016) pp. 873–883. 0031-8248/2016/8305-0020\$10.00
Copyright 2016 by the Philosophy of Science Association. All rights reserved.

This article offers a critical reassessment of the debate over the generality of Hamilton's rule and kin selection theory (see also Birch 2014; Birch and Okasha 2015). My main focus will be on additivity rather than weak selection (although the two issues are linked, since if weak selection is assumed, then additivity of costs and benefits can often be justified as a linear approximation). In brief, I argue that the generalized version of Hamilton's rule favored by Frank (1997, 1998), Gardner et al. (2011), and Marshall (2011), although correct as a statement about allele frequency change, is conceptually impoverished compared to Hamilton's original. This is so for two reasons, both to do with the failure of additivity. The first reason is that without additivity, the generalized form of Hamilton's rule has limited biological meaning, as its components lack a natural causal interpretation. The second reason is that without additivity, the link between Hamilton's rule and the idea that natural selection leads organisms to behave 'as if' trying to maximize their inclusive fitness breaks down.

2. Two Aspects of Kin Selection Theory. As is well known, kin selection theory in the Hamiltonian tradition contains two distinct although related ideas (Okasha and Martens 2016b). The first is Hamilton's rule itself, the rule-of-thumb criterion ($rb > c$) for when an allele that causes a social behavior will be favored by natural selection, where c is the cost to the actor, b the benefit to the recipient, and r the 'coefficient of relationship' between them. The second is the idea that individuals will behave 'as if' trying to maximize their inclusive fitness, a measure that takes into account an individual's contribution to the reproductive success of its relatives, rather than their classical (or 'personal') fitness.¹

These two aspects of kin selection theory have not always been sharply distinguished; as a result, the relation between them is not fully settled. Many expositions of Hamilton's theory (e.g., Bourke 2011; Marshall 2015) focus exclusively on the first aspect. However, recently Grafen (2006, 2009), Gardner et al. (2011), and Queller (2011) have argued for the central importance of the idea of inclusive fitness maximization as the 'goal' of an individual's social behavior. Moreover, Grafen (2006) tries to integrate this idea with an explicit population-genetic description of the evolution of social behavior, which goes some way toward bringing the two aspects of kin selection theory into harmony with each other.

Importantly, the analysis of Grafen (2006) assumes additivity of costs and benefits, as he points out. In the context of the debate over the generality of Hamilton's rule, this prompts an immediate question. If we follow

1. This maximization claim should be distinguished from the claim that selection will act to maximize the average inclusive fitness of the whole population; the latter claim is found in Hamilton (1964), the former in Hamilton (1970).

Gardner et al. (2011) and others who favor a generalized version of Hamilton's rule, which holds true even in the nonadditive case, can we recover the principle of inclusive fitness maximization using an argument similar to that of Grafen (2006)? I argue that the answer is no, for conceptual rather than technical reasons. In particular, if we define the cost and benefit terms the way they must be defined for Hamilton's rule to be a universally correct statement, as partial regression coefficients, then the inclusive fitness that results is not a quantity that individuals can sensibly be regarded as trying to maximize, as it is not solely a function of their own behavior.

3. Nonadditivity: Preliminaries. To focus the issue, consider first a gene for a nonsocial trait that affects individual fitness. Assuming no mutation or gametic selection, the one-generation change in the population-wide frequency of the gene is given by

$$\Delta \bar{p} = \frac{\text{Cov}(w_i, p_i)}{\bar{w}}, \quad (1)$$

where w_i is the fitness (i.e., gametic output) of the i th individual, p_i is the frequency of the gene in the i th individual, \bar{p} is the population-wide frequency of the gene, and \bar{w} is average population fitness (Price 1970). Presuming that $\text{Var}(p_i) > 0$, the covariance term in equation (1) can be decomposed to give

$$\Delta \bar{p} = b_{wp} \frac{\text{Var}(p_i)}{\bar{w}}, \quad (2)$$

where b_{wp} is the linear regression of individual fitness w_i on individual gene frequency p_i . So the condition for the gene to increase is $b_{wp} > 0$. Note that (1) follows from (2) simply by the definition of the linear regression coefficient. Note also that (1) and (2) would both hold true if for p_i and \bar{p} we substituted the 'breeding value' of some phenotypic trait and its population-wide average (Grafen 1985; Falconer 1995).

Equation (2) makes no assumption that the true causal dependence of w_i on p_i is linear. Even if that dependence is highly nonlinear, the slope of the best-fit regression line of fitness against individual gene frequency is what determines whether the gene will spread. Suppose, for example, that the true causal relation is given by $w_i = 1 + \beta p_i^2$, where $\beta > 0$, which means that the fitness difference between individuals with zero and one copies of the gene is less than that between individuals with one and two copies. This has two important implications. First, the linear regression coefficient b_{wp} will be a function of the population's genotypic composition, so it will change as the population evolves. By contrast if w_i depends linearly on p_i , then b_{wp} will be independent of the gene's frequency in the population.

Second, in the nonlinear case b_{wp} cannot be construed as a measure of the strength of the causal influence of the gene on fitness and is not counterfactually informative. That is, b_{wp} does not tell us the fitness difference that an individual would incur if she were to receive an extra copy of the gene. Rather, b_{wp} is a measure of the average fitness difference between individuals with x and $x + 1$ copies of the gene in the actual population. This is an instance of a well-known point in the statistics literature, namely, that a linear regression analysis can only be used to tell us the effect of a hypothetical intervention if the regression model correctly describes the true causal relations in the world (see, e.g., Gelman and Hill 2007, chap. 9). The relevance of this for Hamilton's rule will become clear.

This last point can be made more vivid by noting that equations (1) and (2) hold true even if the gene has no causal effect at all on an individual's fitness but correlates with fitness for some other reason. In such a circumstance, the value of b_{wp} may be substantial, but if an individual in the population were given another copy of the gene, for example, by mutation, her fitness would remain unchanged. As in the nonlinear case, the fact that b_{wp} correctly predicts the evolutionary change, but does not correctly predict the effect of hypothetical alterations to an individual's genotype, reflects the fact that the simple linear regression of w_i on p_i does not correctly model the true causal determinants of individual fitness.

This leads to our final preliminary point. Although we stipulated above that the gene codes for a nonsocial trait, this fact plays no role in the derivation of equation (1), which in fact holds true quite generally. An important consequence of this, emphasized by Grafen (2002), is that one cannot infer from equation (1) that selection will lead individuals to maximize their individual fitness w_i . This is generally true for a nonsocial trait but not for a social trait, yet equation (1) applies in both cases. Grafen (2002) captures this point by distinguishing between a 'target for selection', that is, a quantity that covaries with p_i in equation (1), from an 'individual maximand', that is, a quantity that an individual behaves as if trying to maximize. The point can equally be captured by noting that in the social case, an individual's fitness w_i does not causally depend solely on its own genotype p_i but also on the genotypes of its social partners. So it does not make sense to regard w_i as a quantity that an individual might seek to maximize through its behavioral choices. As before, the linear regression b_{wp} correctly predicts the evolutionary change, but its value does not solely reflect the direct causal influence of p_i on w_i , and this is precisely the reason why individual behavior does not maximize w_i .

4. Hamilton's Rule Generalized. A generalized version of Hamilton's rule is easily derived from equation (1), following Queller (1992a, 1992b), Frank (1997, 1998), and Gardner et al. (2011). Suppose that the gene codes

for a social action, so an individual's fitness w_i depends on both its own gene frequency p_i and the gene frequency of its social partners p'_i . We then write w_i as a linear regression on p_i and p'_i :

$$w_i = \alpha + \beta_{wp.p} p_i + \beta_{wp'.p} p'_i + e_i, \quad (3)$$

where α is 'baseline' fitness; $\beta_{wp.p}$ is the partial regression of individual fitness on individual gene frequency, controlling for social partners' gene frequency; $\beta_{wp'.p}$ is the partial regression of individual fitness on social partners' gene frequency, controlling for individual gene frequency; and e_i is the residual whose variance is to be minimized. Note that equation (3) does not presume that p_i and p'_i make additive contributions to an individual's fitness, although if they do not, our previous caveats about not interpreting regression coefficients as measures of causal influence apply.

Equation (3) can then be substituted into equation (1), which after simplification yields

$$\Delta \bar{p} = (\beta_{wp.p} + \beta_{wp'.p} r) \frac{\text{Var}(p)}{\bar{w}}, \quad (4)$$

where $r = \beta_{p'p} = \text{Cov}(p, p') / \text{Var}(p)$ is the regression of social partners' genotype on individual genotype, which is one standard definition of the coefficient of relatedness. Equation (4) is a version of Hamilton's rule in its 'neighbor-modulated' form, that is, that considers the effects on a focal individual's fitness of the genes (and hence actions) of its social partners.

As Hamilton (1964) first showed, we can instead consider the effects of a focal individual's genes (and hence actions) on the fitness of her social partners, rather than vice versa. The relevant regression coefficient corresponding to this effect is $\beta_{w'p.p'}$, that is, the partial regression of social partners' fitness on an individual's gene frequency, controlling for the social partners' gene frequency. It is well known that under quite general conditions, $\beta_{wp'.p} = \beta_{w'p.p'}$, which means that equation (4) can be rewritten as

$$\Delta \bar{p} = (\beta_{wp.p} + \beta_{w'p.p'} r) \frac{\text{Var}(p)}{\bar{w}}. \quad (5)$$

By labeling $\beta_{wp.p}$ and $\beta_{w'p.p'}$ as ' $-c$ ' and ' b ' respectively, we can see that the condition for the spread of the allele in the population is $rb > c$, which is Hamilton's rule in its traditional guise. This version of Hamilton's rule is employed by Queller (1992a), Frank (1997, 1998), Gardner et al. (2011), and Marshall (2011, 2015).

As Frank (1997) and Gardner et al. (2011) rightly stress, equation (5) is a fully general truth about evolution by natural selection that simply partitions the total evolutionary change into direct and indirect components and is thus as general as the Price equation itself. In particular, neither weak

selection nor additivity of costs and benefits is needed to derive (5), a fact that Gardner et al. (2011) use to rebut the charge of limited generality leveled against kin selection theory by authors such as Nowak et al. (2010, 2011).

5. Three Issues. The crux of the matter here is whether it is legitimate to identify the $-c$ and b terms of Hamilton's rule with the partial regression coefficients $\beta_{wp \cdot p'}$ and $\beta_{wp' \cdot p}$, derived from the multiple regression equation (3). If this identification is granted, then Hamilton's rule does indeed obtain with complete generality, irrespective of additivity or weak selection. Gardner et al. (2011) insist that defining costs and benefits as partial regression coefficients, that is, average effects, is perfectly correct and indeed criticize other authors for not realizing this and adopting 'arbitrary' definitions of the $-c$ and b terms. However, this raises three interpretive issues.

The first concerns causality. Hamilton's original papers deliberately employ causal language to describe the fitness costs and benefits of social actions (Hamilton 1964, 1970). He is explicit that the 'cost' in question is the reduction in personal reproduction caused by a social action (or gene), and similarly for the benefit. But as stressed above, partial regression coefficients can only be construed as causal under certain circumstances; in particular, the regression model must truly describe the causal dependence of the dependent variable on the independent ones. Now if the true dependence of w_i on p_i and p'_i is nonadditive, or if there are other variables that causally affect w_i and are correlated with p_i or p'_i , then the regression coefficients $\beta_{wp \cdot p'}$ and $\beta_{wp' \cdot p}$ cannot be regarded as natural measures of the causal influence of p_i and p'_i on fitness. Indicative of this is that if we take a given individual who does not have the gene in question and so does not perform the social action and ask what the effect on her individual fitness would be if she did perform the social action while holding fixed her social partners, then the answer is not given by $\beta_{wp \cdot p'}$ (see Okasha and Martens 2016a). So while it is possible to derive equation (5) in the nonadditive case, identifying its components with the $-c$ and b terms of Hamilton's rule sacrifices the causal understanding of these terms.²

The second issue concerns biological significance. As Frank (1997, 1998) and Queller (2011) both stress, it is possible to use any 'predictors' as independent variables in a multiple regression analysis. There is nothing sacrosanct about the choice of p_i and p'_i as predictors in equation (3); other predictors such as the product $p_i \cdot p'_i$, or p_i^2 , could easily be added, which would lead the total evolutionary change to be partitioned up differently from

2. Okasha and Martens (2016a) explore a possible way of salvaging the causal interpretation of Hamilton's rule, by drawing on an argument made by Fisher (1941) in relation to his concept of 'average effect of a gene substitution'.

equation (5). Ideally, we want a good partition to have components that are biologically meaningful. Now if w_i depends nonadditively on p_i and p'_i , then arguably equation (5) fails this criterion. Indicative of this is that the values of $\beta_{wp.p'}$ and $\beta_{w'p.p'}$ then become functions of the population's genotypic composition, as discussed previously. From this perspective, it seems that insisting on applying equation (5) even in nonadditive situations, and preserving the generality of Hamilton's rule by identifying the $-c$ and b terms with the $\beta_{wp.p'}$ and $\beta_{w'p.p'}$ coefficients, comes at the expense of its biological significance (see Allen et al. 2013; Birch 2014; Birch and Okasha 2015, for further discussion).

The third issue is the most important and concerns inclusive fitness maximization. As discussed above, kin selection theory contains two distinct ideas: the $rb > c$ criterion for the spread of a gene for social actions and inclusive fitness maximization as the 'goal' of individual behavior. Even if we are happy to identify the $\beta_{wp.p'}$ and $\beta_{w'p.p'}$ coefficients of equation (5) with the $-c$ and b terms of Hamilton's rule, this only speaks to the first part of kin selection theory; that is, it shows that the $rb > c$ criterion is universally valid. But does it also allow us to recover maximization of inclusive fitness as a universally valid principle? This question is not directly addressed by the authors who favor the generalized form of Hamilton's rule; however, Gardner et al. strongly imply that the answer is yes; they write of equation (5) that "the partition of natural selection into direct and indirect components will exactly correspond with the direct and indirect components of inclusive fitness; the quantity that organisms are designed to maximize (Hamilton 1964, 1996; Grafen 2006) which better clarifies the link between the process and purpose of Darwinian adaptation" (Gardner et al. 2011, 1032).

However, this is a questionable claim. The most careful elaboration of the connection between gene frequency change and inclusive fitness maximization is due to Grafen (2006), whose analysis is explicitly restricted to the case in which costs and benefits are additive. One might hope that this restriction could be lifted with further work, that is, that an argument akin to Grafen's could be developed for the more general nonadditive case, which is presumably what Gardner et al. (2011) believe. But this suggestion faces an immediate conceptual problem, given that the $\beta_{wp.p'}$ and $\beta_{w'p.p'}$ terms in equation (5) are functions of the population's genotypic composition in the nonadditive case. To see the problem, we need to briefly expound the logical structure of Grafen's (2006) argument.

6. Grafen's Defense of Inclusive Fitness Maximization. Grafen's approach is to use a fully explicit definition of optimization and then to seek links between the optimality of individual behavior, in social contexts, and population genetics. The notion of optimization is captured by an 'objective

function' that maps an individual's phenotype (or behavior) to the real numbers; if individuals achieve the maximum value of this function they are said to 'behave optimally'. What Grafen then shows is that if the objective function is taken to be inclusive fitness, defined as $(rb - c)$, then certain logical links between optimality and gene frequency change will hold. In effect, these links say that if and only if every individual chooses a behavior that maximizes her inclusive fitness, then population-genetic equilibrium will result; that is, no gene frequency change will occur and no mutants can invade. Grafen (2006) argues that this constitutes a formal vindication of Hamilton's idea that natural selection on genes for social actions will lead individuals to behave as if maximizing their inclusive fitness.³

From this brief description of Grafen's argument, one point should already be clear. It is essential that the 'objective function', whatever it is taken to be, should be a function solely of an individual's choice of behavior. This is because the point of the argument is to represent individuals as akin to rational agents, choosing between alternative behaviors according to how well they score on some criterion, like the utility maximizers of economic theory. When the objective function is taken to be inclusive fitness, this requirement is satisfied, given that Grafen defines the costs and benefits in Hamilton's original way, as fitness increments caused by individuals' social actions. Indeed Grafen explicitly incorporates an assumption that he calls 'actor's control', which says that the individual actor "controls both the performance of the action and its quantitative consequences" (2006, 553); this implies that the benefit b_{ij} conferred by individual i on individual j "will depend only on the phenotype of individual i " (554). It is precisely because of this that it makes sense to regard individuals as trying to maximize their inclusive fitness (i.e., $rb - c$), as the amount of inclusive fitness that an individual obtains depends solely on what behavior it performs.

It should now be clear why Grafen's argument does not readily generalize to the nonadditive case, where the costs and benefits are defined as the partial regression coefficients of equation (5). For if inclusive fitness is defined using these costs and benefits, the assumption of 'actor's control' immediately fails, since the values of $\beta_{wp.p'}$ and $\beta_{w'p.p'}$ depend on the population's genetic composition. Therefore, the amount of inclusive fitness that an individual gets will not depend solely on its choice of social action; it will also depend on population-wide gene frequencies. So conceptually, it does not make sense to treat $(\beta_{wp.p'} + \beta_{w'p.p'}r)$ as a quantity that an individual might seek to maximize through its choice of action. Thus, when Gardner et al. suggest that "we can imagine the individual adjusting her inclusive

3. For discussion of the logic of Grafen's argument, see the symposium on Grafen's work in the special edition of *Biology and Philosophy* 29, no. 2 (2015).

fitness . . . by altering her behaviour" (2011, 1039–40), in an allusion to Grafen's argument, they are illicitly generalizing from the additive case, in which inclusive fitness is solely a function of individual behavior, to the general case in which it is not. There is a principled reason why a Grafen-style argument for inclusive fitness maximization cannot apply, if the costs and benefits are defined as partial regression coefficients in the manner of equation (5).

The point can be seen from another perspective, by returning to Grafen's distinction between 'target of selection' and 'individual maximand'. In the example of equation (1) applied to a gene coding for a social trait, we saw that although b_{wp} is the correct criterion for the spread of the gene, it is not true that individual fitness w_i is the maximand of an individual's behavior; it cannot be since an individual's w_i does not depend solely on its own behavior. A similar moral applies in the case of equation (5), the generalized Hamilton's rule, applied to the case of nonadditive costs and benefits. The criterion for the spread of the gene is $(\beta_{wp.p'} + \beta_{w'p.p'}r) > 0$, but it is not true that individuals will behave as if maximizing $(\beta_{wp.p'} + \beta_{w'p.p'}r)$; they cannot do this, since the value of that quantity that an individual receives does not solely depend on its own behavior. In both cases, the target of selection cannot be equated with the maximand of individual behavior, for the same reason. This highlights the crucial importance of the notion of actor's control for arguments about optimization.

The point can also be related to our foregoing observations about causality. A rational agent seeking to maximize its inclusive fitness, or an organism modeled as such an agent, needs to be able consider the consequences for its inclusive fitness of different possible actions. Thus, they must entertain conditional statements such as "if I were to perform action x , I would suffer cost c , but reap indirect benefits rb ." But as we have seen, in the general case the partial regression coefficients $\beta_{wp.p'}$ and $\beta_{w'p.p'}$ cannot be equated with the costs and (indirect) benefits that an individual would have obtained by hypothetically altering its value of p_i (i.e., its action) while keeping p'_i fixed; such an interpretation is only permissible if the true causal dependence of w_i on p_i and p'_i is additive. So while it is true that in the nonadditive case, the average effects $\beta_{wp.p'}$ and $\beta_{w'p.p'}$ are what matter to natural selection, in that they determine the gene frequency change, they are not what matter to an individual agent deliberating about what to do.

Is there any way to salvage the idea of inclusive fitness maximization for the case of nonadditive payoffs? The fact that Grafen's own argument requires additivity, because of his assumption of actor's control, does not preclude some other argument for the desired conclusion being given. Okasha and Martens (2016b) develop a different way of formalizing the idea that evolution will lead individuals to engage in maximizing behavior, in relation to symmetric two-player games; the idea is simply to find a utility func-

tion such that when both players play Nash equilibrium strategies, evolutionary equilibrium obtains and vice versa. Unlike Grafen's approach, this approach can in principle handle nonadditive payoffs, since instead of simple optimization it substitutes 'best response', that is, optimal behavior conditional on a partner's behavior. However, it does not vindicate inclusive fitness maximization. By considering a simple nonadditive prisoner's dilemma model of social evolution (or 'synergy game'), Okasha and Martens (2016b) find that at evolutionary equilibrium, the relevant utility function, which individuals behave as if they are trying to maximize, is not the inclusive fitness function. This result suggests that recovering inclusive fitness maximization for nonadditive payoffs is unlikely to work.

7. Conclusion. The view that Hamilton's rule represents a fully general statement about natural selection, which holds true irrespective of whether costs and benefits are additive, is an interesting one. Recent work has shown clearly that this view is correct, so long as the $-c$ and b terms are appropriately defined as average effects, that is, partial regression coefficients, as in equation (5). However, when costs and benefits are defined this way, the causal meaning and biological significance of the rule are arguably sacrificed, and moreover, the 'inclusive fitness' that results is not a quantity that individual organisms can sensibly be thought of as trying to maximize. So while the first aspect of kin selection theory (the $rb > c$ criterion for allele frequency change) can be salvaged as a fully general principle, applicable even if costs and benefits are nonadditive, the second aspect of the theory (individuals behaving as if maximizing inclusive fitness) is a rather different matter. While there may be a way of extending this second idea to the non-additive case, it is unclear how this can be done; Grafen's (2006) argument will not work, since it requires that an individual's objection function depend only on her choice of action, and the game-theoretic approach developed by Okasha and Martens (2016b) suggests that inclusive fitness is not in fact the quantity that individuals will behave as if they want to maximize, in a simple nonadditive model.

REFERENCES

- Allen, Benjamin, Martin A. Nowak, and Edward O. Wilson. 2013. "Limitations of Inclusive Fitness." *Proceedings of the National Academy of Sciences USA* 110:21035–39.
- Birch, Jonathan. 2014. "Hamilton's Rule and Its Discontents." *British Journal for the Philosophy of Science* 65:381–411.
- Birch, Jonathan, and Samir Okasha. 2015. "Kin Selection and Its Critics." *Bioscience* 65:22–32.
- Bourke, Andrew F. G. 2011. *Principles of Social Evolution*. Oxford: Oxford University Press.
- Falconer, Douglas S. 1995. *Introduction to Quantitative Genetics*. 4th ed. London: Longman.
- Fisher, Ronald A. 1941. "Average Effect and Average Excess of a Gene Substitution." *Annals of Human Genetics* 11:53–63.

- Frank, Steven A. 1997. "The Price Equation, Fisher's Fundamental Theorem, Kin Selection, and Causal Analysis." *Evolution* 51:1712–29.
- . 1998. *Foundations of Social Evolution*. Princeton, NJ: Princeton University Press.
- Gardner, Andy, Stuart A. West, and Geoff Wild. 2011. "The Genetical Theory of Kin Selection." *Journal of Evolutionary Biology* 24:1020–43.
- Gelman, Andrew, and Jennifer Hill. 2007. *Data Analysis Using Regression and Multi-level/Hierarchical Models*. Cambridge: Cambridge University Press.
- Grafen, Alan. 1985. "A Geometric View of Relatedness." *Oxford Surveys in Evolutionary Biology* 2:28–89.
- . 2002. "A First Formal Link between the Price Equation and an Optimization Program." *Journal of Theoretical Biology* 217:75–91.
- . 2006. "Optimization of Inclusive Fitness." *Journal of Theoretical Biology* 238:541–63.
- . 2009. "Formalizing Darwinism and Inclusive Fitness Theory." *Philosophical Transactions of the Royal Society B* 364:3135–41.
- Hamilton, William D. 1964. "The Genetical Evolution of Social Behaviour." Pts. 1 and 2. *Journal of Theoretical Biology* 7:1–52.
- . 1970. "Selfish and Spiteful Behaviour in an Evolutionary Model." *Nature* 228:1218–20.
- . 1996. *Narrow Roads of Gene Land*. Vol. 1. New York: Freeman.
- Marshall, James A. R. 2011. "Group Selection and Kin Selection: Formally Equivalent Approaches." *Trends in Ecology and Evolution* 26:325–32.
- . 2015. *Social Evolution and Inclusive Fitness Theory*. Princeton, NJ: Princeton University Press.
- Nowak, Martin A., Corina E. Tarnita, and Edward O. Wilson. 2010. "The Evolution of Eusociality." *Nature* 466:1057–62.
- . 2011. "Nowak et al. Reply." *Nature* 471:E0–E10.
- Okasha, Samir, and Johannes Martens. 2016a. "The Causal Meaning of Hamilton's Rule." *Royal Society Open Science*. doi:10.1098/rsos.160037.
- . 2016b. "Hamilton's Rule, Inclusive Fitness Maximization, and the Goal of Individual Behaviour in Symmetric Two-Player Games." *Journal of Evolutionary Biology* 29:473–82.
- Price, George R. 1970. "Selection and Covariance." *Nature* 227:520–21.
- Queller, David C. 1992a. "A General Model for Kin Selection." *Evolution* 46:376–80.
- . 1992b. "Quantitative Genetics, Inclusive Fitness and Group Selection." *American Naturalist* 139:540–58.
- . 2011. "Expanded Social Fitness and Hamilton's Rule for Kin, Kith and Kind." *Proceedings of the National Academy of Sciences USA* 108:10792–99.
- van Veelen, Matthias. 2009. "Group Selection, Kin Selection, Altruism and Cooperation: When Inclusive Fitness Is Right and When It Can Be Wrong." *Journal of Theoretical Biology* 259: 589–600.